A. Harris

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/397,558*

DATE: 10/21/1999
TIME: 09:32:35

INPUT SET: S33714.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

#7

ENTERED

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Lal, Preeti
Guegler, Karl J.
Corley, Neil C.

(ii) TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
(B) STREET: 3174 PORTER DRIVE
(C) CITY: PALO ALTO
(D) STATE: CALIFORNIA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/397,558
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 09/083,521
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: CERRONE, MICHAEL C.
(B) REGISTRATION NUMBER: 39,132
(C) REFERENCE/DOCKET NUMBER: PF-0527 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (650) 855-0555
(B) TELEFAX: (650) 845-4166

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: PROSTUT10
(B) CLONE: 1691243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1 :

```
Met Val His Val Ala Tyr Ser Leu Cys Leu Pro Met Arg Arg Ser
                  5                  10                  15
Glu Arg Tyr Leu Phe Leu Asn Met Ala Tyr Gln Gln Val His Ala
             20                  25                  30
Asn Ile Glu Asn Ser Trp Asn Glu Glu Glu Val Trp Arg Ile Glu
         35                  40                  45
Met Tyr Ile Ser Phe Gly Ile Met Ser Leu Gly Leu Leu Ser Leu
     50                  55                  60
Leu Ala Val Thr Ser Ile Pro Ser Val Ser Asn Ala Leu Asn Trp
 65                  70                  75
Arg Glu Phe Ser Phe Ile Gln Ser Thr Leu Gly Tyr Val Ala Leu
             80                  85                  90
Leu Ile Ser Thr Phe His Val Leu Ile Tyr Gly Trp Lys Arg Ala
         95                 100                 105
Phe Glu Glu Glu Tyr Tyr Arg Phe Tyr Thr Pro Pro Asn Phe Val
    110                 115                 120
Leu Ala Leu Val Leu Pro Ser Ile Val Ile Leu Asp Leu Leu Gln
125                 130                 135
Leu Cys Arg Tyr Pro Asp
140
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 410 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: BRSTTUT03

(B) CLONE: 1999442

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2 :

```
Met Phe Leu Pro Pro Val Val Leu Ala Ile Arg Ser Arg Tyr Val
                  5                  10                  15
Leu Glu Ala Ala Val Tyr Thr Phe Thr Met Phe Phe Ser Thr Phe
                 20                  25                  30
Tyr His Ala Cys Asp Gln Pro Gly Ile Val Val Phe Cys Ile Met
                 35                  40                  45
Asp Tyr Asp Val Leu Gln Phe Cys Asp Phe Leu Gly Ser Leu Met
                 50                  55                  60
Ser Val Trp Val Thr Val Ile Ala Met Ala Arg Leu Gln Pro Val
                 65                  70                  75
Val Lys Gln Val Leu Tyr Leu Leu Gly Ala Met Leu Leu Ser Met
                 80                  85                  90
Ala Leu Gln Leu Asp Arg His Gly Leu Trp Asn Leu Leu Gly Pro
                 95                 100                 105
Ser Leu Phe Ala Leu Gly Ile Leu Ala Thr Ala Trp Thr Val Arg
                110                 115                 120
Ser Val Arg Arg Arg His Cys Tyr Pro Pro Thr Trp Arg Arg Trp
                125                 130                 135
Leu Phe Tyr Leu Cys Pro Gly Ser Leu Ile Ala Gly Ser Ala Val
                140                 145                 150
Leu Leu Tyr Ala Phe Val Glu Thr Arg Asp Asn Tyr Phe Tyr Ile
                155                 160                 165
His Ser Ile Trp His Met Leu Ile Ala Gly Ser Val Gly Phe Leu
                170                 175                 180
Leu Pro Pro Arg Ala Lys Thr Asp His Gly Val Pro Ser Gly Ala
                185                 190                 195
Arg Ala Arg Gly Cys Gly Tyr Gln Leu Cys Ile Asn Glu Gln Glu
                200                 205                 210
Glu Pro Gly Pro Arg Gly Pro Arg Arg Gly His Cys Gln Gln His
                215                 220                 225
Leu Cys Gln Leu Arg Gly Ala Leu Gly Leu Ala Leu Arg Gly Tyr
                230                 235                 240
Glu Cys Phe Leu Glu Phe Phe Leu Gly Val Trp Ser Pro Leu Arg
                245                 250                 255
Arg Arg Gln Ala Val Phe Leu Glu Asp Met Glu Ser Phe Ser Arg
                260                 265                 270
Thr Gln Asn Ser Ser Arg Asp Leu Glu Pro Phe Pro Gly His Gly
                275                 280                 285
Glu Leu Pro Glu Gly Leu Glu Ser Pro Cys Ile Met Glu Ser Phe
                290                 295                 300
Leu Arg Thr Gly Ala Tyr Ala Gly Thr Glu Ser Leu Arg Thr Lys
                305                 310                 315
Glu Ser Leu Leu Gln Val Trp Ser Leu Ser Trp Asp Ala Glu Pro
                320                 325                 330
Ser Gln Asp Met Asp Ser Phe Pro Gly Arg Gln Ser Pro Val Arg
                335                 340                 345
Ser Thr Ala Ser Phe Gln Arg Arg Trp Ser Leu Ser Trp Gly Asn
                350                 355                 360
Gln Ile Ser Arg Phe Ser Gln Arg Leu Ser Asn Ser Gly Leu Arg
```

365 370 375
Leu Pro Ser Gln Arg Gln Arg Leu Gly Cys Ala Val Leu Trp Arg
380 385 390
Arg Asp Cys Arg Met Asp Gly Ala Gly Thr Gly Ala Val Trp Val
395 400 405
Ala Gly Ile Leu Val
410

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: PROSTUT10
(B) CLONE: 1691243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3 :

CAAGTATAGG AGATTTCCAC CTTGGTTGGA AACCTGGTTA CAGTGTAGAA AACAGCTTGG 60
ATTACTAAGT TTTTTCTTCG CTATGGTCCA TGTTGCCTAC AGCCTCTGCT TACCGATGAG 120
AAGGTCAGAG AGATATTTGT TTCTCAACAT GGCTTATCAG CAGGTTCATG CAAATATTGA 180
AAACTCTTGG AATGAGGAAG AAGTTTGGAG AATTGAAATG TATATCTCCT TTGGCATAAT 240
GAGCCTTGGC TTACTTTCCC TCCTGGCAGT CACTTCTATC CCTTCAGTGA GCAATGCTTT 300
AAACTGGAGA GAATTCAGTT TTATTCAGTC TACACTTGGA TATGTCGCTC TGCTCATAAG 360
TACTTTCCAT GTTTTAATTT ATGGATGGAA ACGAGCTTTT GAGGAAGAGT ACTACAGATT 420
TTATACACCA CCAAACTTTG TTCTTGCTCT TGTTTTGCCC TCAATTGTAA TTCTGGATCT 480
TTTGCAGCTT TGCAGATACC CAGACTGAGC TGGAACTGGA ATTTGTCTTC CTATTGACTC 540
TACTTCTTTA AAAGCGGCTG CCCATTACAT TCCTCAGCTG TCCTTGCAGT TAGGTGTACA 600
TGTGACTGAG TGTTGGCCAG TGAGATGAAG TCTCCTCAAA GGAAGGCAGC ATGTGTCCTT 660
TTTCATCCCT TCATCTTGCT GCTGGGATTG TGGATATAAC AGGAGCCCTG GCAGCTGTCT 720
CCAGAGGATC AAAGCCACAC CCAAAGAGTA AGGCAGATTA GAGACCAGAA AGACCTTGAC 780
TACTTCCCTA CTTCCACTGC TTTTTCCTGC ATTTAAGCCA TTGTAAATCT GGGTGTGTTA 840
CATGAAGTGA AAATTAATTC TTTCTGCCCT TCAGTTCTTT ATCCTGATAC CATTTAACAC 900
TGTCTGAATT AACTAGACTG CAATAATTCT TTCTTTTGAA AGCTTTTAAA GGATAATGTG 960
CAATTCACAT TAAAATTGAT TTTCCATTGT CAATTAGTTA TACTCATTTT CCTGCCTTGA 1020
TCTTTCATTA GATATTTTGT ATCTGCTTGG AATATATTAT CTTCTTTTTA ACTGTGTAAT 1080
TGGTAATTAC TAAAACTCTG TAATCTCCAA AATATTGCTA TCAAATTACA CACCATGTTT 1140
TCTATCATTC TCATAGATCT GCCTTATAAA CATTTAAATA AAAAGTACTA TTTAATGATT 1200
TAAAAAAAAA AAA 1213

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1297 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: BRSTTUT03
(B) CLONE: 1999442

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4 :

```
CGGACGCGTG GGCTGCTCTG CCTGAGCAAC CTCATGTTTC TGCCACCTGT GGTCCTGGCC  60
ATTCGGAGTC GATATGTGCT GGAAGCTGCA GTCTACACCT TCACCATGTT CTTCTCCACG 120
TTCTATCATG CCTGTGACCA GCCAGGCATC GTGGTTTTCT GCATCATGGA CTACGATGTG 180
CTGCAGTTCT GTGATTTCCT GGGCTCCTTA ATGTCCGTGT GGGTCACTGT CATTGCCATG 240
GCTCGTTTAC AGCCCGTGGT CAAGCAGGTG CTGTATTTGC TGGGAGCTAT GCTGCTGTCC 300
ATGGCTCTGC AGCTTGACCG ACATGGACTC TGGAACCTGC TTGGACCCAG TCTCTTCGCC 360
CTGGGGATCT TGGCCACAGC CTGGACAGTA CGCAGCGTCC GCCGCCGGCA CTGCTACCCA 420
CCCACGTGGC GCCGCTGGCT TTTCTACTTG TGCCCTGGCA GCCTTATTGC AGGCAGTGCC 480
GTCCTGCTTT ATGCTTTTGT GGAGACCCGG GACAACTACT TCTACATTCA CAGCATTTGG 540
CATATGCTCA TTGCGGGCAG TGTGGGCTTC CTGCTGCCCC CTCGTGCCAA GACTGACCAC 600
GGGGTCCCAT CTGGAGCCCG GGCCCGGGGC TGTGGTTACC AGCTATGCAT CAACGAGCAG 660
GAGGAGCCTG GGCCTCGTGG GCCCAGGAGG GGCCACTGTC AGCAGCATCT GTGCCAGCTG 720
AGAGGGGCTT TGGGCCTGGC CCTGAGGGGA TATGAATGCT TCCTAGAGTT CTTTCTGGGG 780
GTGTGGAGCC CTCTTAGAAG GAGACAGGCT GTATTTCTTG AGGACATGGA GTCTTTCTCA 840
AGGACACAAA ACTCTTCCAG GGACCTGGAG CCCTTCCCAG GACATGGAGA ACTTCCTGAG 900
GGCCTGGAGT CCCCCTGCAT CATGGAGTCC TTCTTAAGGA CTGGAGCCTA TGCAGGCACA 960
GAGTCCCTCA GGACCAAGGA GTCCCTCCTG CAGGTGTGGA GCCTTTCCTG GGATGCAGAG 1020
CCTTCCCAAG ACATGGATTC CTTCCCAGGG AGACAAAGCC CTGTCAGGAG CACAGCATCT 1080
TTCCAGAGGA GGTGGAGTCT ATCTTGGGGA AACCAAATTT CCAGATTTTC CCAGAGGCTC 1140
AGCAACTCTG GCCTCAGGCT TCCTTCCCAG AGGCAGCGTC TGGGCTGTGC TGTGCTGTGG 1200
AGGAGGGATT GCAGGATGGA TGGAGCTGGG ACTGGGGCTG TCTGGGTGGC TGGTATCCTC 1260
GTTTGATACA GGTGGAGTCT CTGTGTCTCC ATAGAAG                           1297
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 1216498

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5 :

```
Met Gly Arg Ala Met Val Val Arg Leu Gly Leu Gly Leu Leu Leu
 1               5                  10                  15
Leu Ala Leu Leu Leu Pro Thr Gln Ile Tyr Cys Asn Gln Thr Ser
            20                  25                  30
Val Ala Pro Phe Ser Gly Asn Gln Ser Ile Ser Ala Ala Pro Asn
        35                  40                  45
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/09/397,558*

INPUT SET: S33714.raw

Line	Error	Original Text

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION *US/09/397,558*

DATE: 10/21/1999
TIME: 09:32:37

INPUT SET: S33714.raw

< < THERE ARE NO ITEMS MISSING > >

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/09/397,558*

DATE: 10/21/1999
TIME: 09:32:37

INPUT SET: S33714.raw

Line	Original Text	Corrected Text